A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design

Graeme L Card[1], Landy Blasdel[1,4], Bruce P England[1], Chao Zhang[1], Yoshihisa Suzuki[1], Sam Gillette[1], Daniel Fong[1], Prabha N Ibrahim[1], Dean R Artis[1], Gideon Bollag[1], Michael V Milburn[1,4], Sung-Hou Kim[2], Joseph Schlessinger[3] & Kam Y J Zhang[1]

Cyclic nucleotide phosphodiesterases (PDEs) comprise a large family of enzymes that regulate a variety of cellular processes. We describe a family of potent PDE4 inhibitors discovered using an efficient method for scaffold-based drug design. This method involves an iterative approach starting with low-affinity screening of compounds followed by high-throughput cocrystallography to reveal the molecular basis underlying the activity of the newly identified compounds. Through detailed structural analysis of the interaction of the initially discovered pyrazole carboxylic ester scaffold with PDE4D using X-ray crystallography, we identified three sites of chemical substitution and designed small selective libraries of scaffold derivatives with modifications at these sites. A 4,000-fold increase in the potency of this PDE4 inhibitor was achieved after only two rounds of chemical synthesis and the structural analysis of seven pyrazole derivatives bound to PDE4B or PDE4D, revealing the robustness of this approach for identifying new inhibitors that can be further developed into drug candidates.

The prevailing approach for drug discovery is to identify potent compounds for a specific target from a large library (500,000–2,000,000) of drug-like chemicals using high-throughput screening. Generally, the initial data generated are the percentage inhibition of all compounds at a single concentration (usually <10 μM); the 50% inhibition concentration (IC_{50}) is then determined for selected compounds, typically those with >80% inhibition at the initial concentration. Those compounds with the highest potency toward the specific molecular target are then chemically optimized. This method has been used to identify a great variety of inhibitors, and recent advances in combinatorial chemistry have increased library sizes to over a million compounds. However, even the largest available compound collections that can be used for screening represent only a small fraction of the vast and diverse repertoire of potential chemicals. In addition, the most potent compounds identified during initial screening are not necessarily the most suitable for further chemical optimization because of undesirable chemical properties or poor drug-like qualities[1].

Alternative approaches that could enable the discovery of compounds from unexploited chemical space have been proposed[2,3]. They are designed to screen, by biophysical methods such as NMR or X-ray crystallography, a small collection of basic chemical building blocks, termed 'fragments', with a molecular weight (MW) generally less than 150 Da, and then to enlarge or link them to increase potency and selectivity[4–6]. However, the difficulty of identifying weakly binding fragments and elaborating or linking them into high-affinity compounds remains a formidable challenge[2].

We have developed an efficient and general approach for the identification of weakly active, low-molecular-weight inhibitors that can be further developed into drug candidates. In this report, we describe how this approach was applied to discover a class of PDE inhibitors.

PDEs are enzymes that play critical roles in maintaining cellular levels of cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP)[7–9], ubiquitous second messengers that mediate numerous biological responses to a great variety of extracellular stimuli. Increased concentrations of cAMP or cGMP result in the activation of cAMP- or cGMP-dependent protein kinases that phosphorylate other signaling proteins and transcription factors to regulate a myriad of physiological processes[10]. Cellular levels of cAMP and cGMP are determined by the relative activities of adenylyl and guanylyl cyclases, which synthesize cyclic nucleotides, and PDEs, which hydrolyze them into 5′-nucleotide monophosphates. PDE inhibition results in higher levels of cyclic nucleotides. Therefore, PDE inhibitors have considerable therapeutic utility as anti-inflammatory agents, anti-asthmatics, vasodilators, smooth muscle relaxants, cardiotonic agents, antidepressants, antithrombotics and agents for improving cognitive functions[11–16].

There are 11 subfamilies of human PDEs, and the cAMP-selective PDE4 subfamily comprises four members[17,18]. PDE4B is of particular importance in the inflammatory responses of lymphocytes. Genetic knockout of PDE4B results in viable mice, yet monocytes from these mice exhibit sharply reduced cytokine production in response to

[1]Plexxikon Inc., 91 Bolivar Dr., Berkeley, California 94710, USA. [2]Department of Chemistry, University of California, Berkeley, California 94720, USA. [3]Department of Pharmacology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA. [4]Present addresses: Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA (L.B.) and Sirtris Pharmaceuticals, 100 Beaver St., Suite 240, Waltham, Massachusetts 02453, USA (M.V.M.). Correspondence should be addressed to K.Y.J.Z. (kzhang@plexxikon.com).
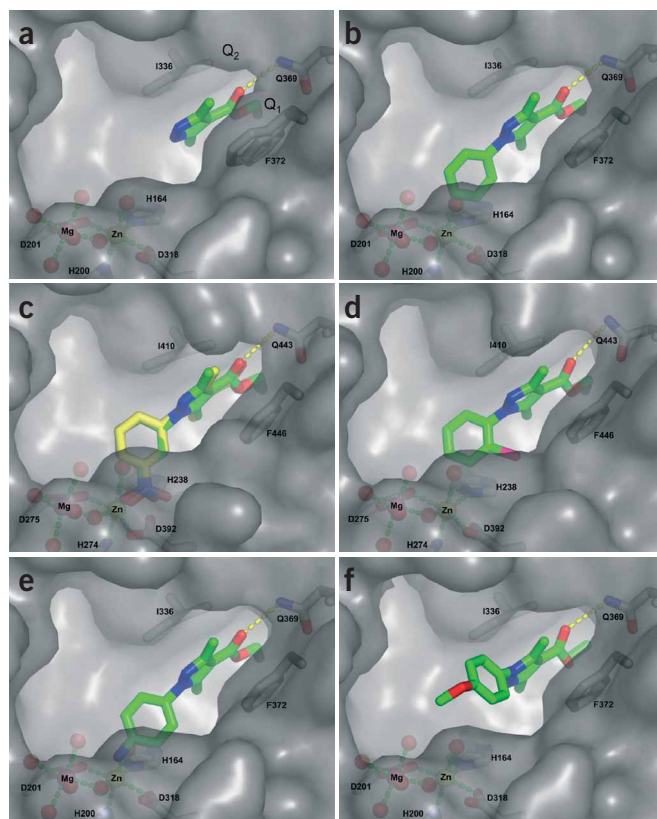
Figure 1 Crystal structure of the pyrazole scaffold and its derivatives in complex with PDE4B or PDE4D. (**a**) Crystal structure of 3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 2) bound to PDE4D, showing the pyrazole ring sandwiched in the hydrophobic clamp formed by F372 and I336. The conserved H-bond, seen in all pyrazole derivative cocrystal structures, between the NE2 atom of the invariant glutamine and the carboxylate group, is shown. (**b**) The crystal structure of 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 8) bound to PDE4D, showing the same interactions as its parent compound, and thus validating the dimethyl pyrazole as a scaffold. The dimethyl pyrazole is sandwiched by F372 and I336 and the carbonyl oxygen forms an H-bond with Q369. The ethoxy group is tucked into the Q1 pocket. (**c**) Crystal structure of 3,5-dimethyl-1-(3-nitro-phenyl)-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 21) bound to PDE4B and PDE4D. The carbon atoms of pyrazole no. 21 bound to PDE4B and PDE4D are shown in green and yellow respectively. The NO_2 group at the meta-position of the phenyl ring formed H-bonds with T345, D392 in PDE4B and the two water molecules coordinating Zn^{2+} (omitted for clarity). (**d**) Crystal structure of 1-(2-chloro-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 20) bound to PDE4B. The Cl-substitution at the ortho-position of the phenyl ring makes several hydrophobic contacts with residues M347, L393 and F446. (**e**) Crystal structure of 1-(4-amino-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 19) bound to PDE4D. The amine group forms three H-bonds with three water molecules, two of which are coordinated to Mg^{2+}. However, this amine nitrogen is also in close proximity to the carbon atom in M273 which results in unfavorable interactions. (**f**) Crystal structure of 1-(4-methoxy-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 17) bound to PDE4D. The methoxy-phenyl group rotated 180° to point away from the di-metal ions to avoid the repulsive interactions between the methoxy group and the di-metal ions.

lipopolysaccharide challenge[19]. One of the shortcomings of the PDE4 inhibitors currently in clinical development is a narrow therapeutic window between efficacy and side effects (e.g., emesis). The side effects may be associated with the specific chemotypes, and there is therefore an urgent need to identify new classes of PDE4 inhibitors that exhibit less emetic effect at the effective therapeutic dose.

The method for scaffold-based drug design presented here uses a starting library of about 20,000 compounds with molecular weights between 125 and 350 Da and a combination of biochemical assays and high-throughput cocrystallography for the primary screening. The central difference between the scaffold-based approach and traditional high-throughput screening methods is that in the former, the typical compounds identified in the initial screen are 100- to 500-fold less potent than in the latter, and the initial synthetic effort is guided by the cocrystal structure of the low-affinity compound in complex with the target protein. We describe the discovery of a class of PDE4 inhibitors using a scaffold identified by this process through the synthesis and screening of a relatively small number of compounds.

RESULTS

To overcome the limited repertoire of compounds that can serve as scaffolds for drug discovery, we developed an efficient strategy for identifying new families of chemical scaffolds that specifically bind to and either inhibit or regulate the activity of given target proteins. This approach enables the discovery of scaffolds that would have been missed by traditional high-throughput screening methods because of their low affinity. Furthermore, it allows compounds to be optimized by adding the least amount of molecular weight to reach a desired potency, thus improving the chances of achieving drug-like properties. Our approach to scaffold-based drug discovery

consists of three steps: scaffold discovery, scaffold validation and chemical optimization.

The first step: scaffold discovery

Scaffold candidates are identified using low-affinity screening of a low-molecular-weight compound library and high-throughput cocrystallography to reveal the molecular basis underlying the activity of the newly identified compounds. To maximize the chance of discovering novel scaffolds, we designed a core compound library by selecting a nonredundant and diverse set of low-molecular-weight compounds from commercially available chemicals. These scaffolds can be further developed into more potent and selective inhibitors by an iterative process involving cocrystallography of a desired compound in complex with a target protein, followed by structure guided computational design of new compounds and subsequent chemical syntheses of these compounds.

We first screened the core library of about 20,000 compounds against a representative subset of PDE family members, PDE1B, PDE2A, PDE4D, PDE5A, PDE7B, using a high-throughput scintillation proximity assay[20]. A total of 316 compounds showed >30% inhibition at 200 μM for three or more PDEs in the screening panel. The 316 compounds were set up for cocrystallization with PDE4D and PDE4B; of these, 269 compounds were cocrystallized and 107 cocrystal structures were solved. A low-affinity 3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (PCEE, **Supplementary Table 1** online, pyrazole no. 2, MW = 168 Da, IC_{50} = 82 μM for PDE4D) revealed the characteristic features of a potential scaffold binding to PDEs[21]: the pyrazole ring is sandwiched in between the hydrophobic clamp formed by residues F372[4D] and I336[4D] (where the superscript identifies the protein to which the residue number refers); the carboxylate of the pyrazole is hydrogen bonded to the invariant purine-selective Q369[4D] (**Fig. 1a**).

Table 1 Structure-activity relationship (SAR) of substitutions on the carboxyl pyrazole scaffold
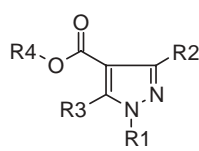
R1	R2	R3	R4	IC_{50} (μM; PDE4B)	IC_{50} (μM; PDE4D)
H-	CH_3-	CH_3-	H-	>200	>200
H-	CH_3-	CH_3-	CH_3-CH_2-	60	82
H-	H-	CH_3-	CH_3-CH_2-	>200	>200
H-	H-	NH_2-	CH_3-CH_2-	>200	>200
H-	H-	CF_3-	CH_3-CH_2-	>200	>200
H-	CH_3-	MXZ-	CH_3-CH_2-	>200	>200
H-	CH_3-	4-CH_3-Ph-	CH_3-CH_2-	15	19
Ph-	CH_3-	CH_3-	CH_3-CH_2-	0.31	0.27
Ph-	4-Cl-Ph-	H-	CH_3-CH_2-	1.5	0.88
QLY-	CH_3-	CH_3-	CH_3-CH_2-	17	19
BTP-	H-	NH_2-	CH_3-CH_2-	25	50

IC_{50} values are in μM.

1. MXZ-: 5-methyl-isoxazolyl
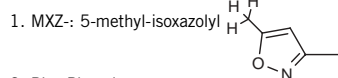
2. Ph-: Phenyl

3. QLY-: 8-quinolinyl
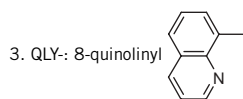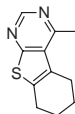
4. BTP-: 4-methyl-5,6,7,8-tetrahydro-benzo[4,5]thieno[2,3-d]pyrimidinyl

The second step: scaffold validation

In the scaffold validation step, we examine whether the newly identified compound can serve as a scaffold for further chemical optimization. A scaffold should not only form key interactions with the protein target[21] but also possess a conserved binding mode that is tolerant to small substitutions. For this purpose, a small number of derivatives of the scaffold candidate were synthesized, and the nature of their interactions with the target protein was analyzed by cocrystallography. A scaffold is considered validated only when the conserved portion of the derivative compounds makes the same interactions with the target protein as the scaffold itself.

To determine whether PCEE could serve as a scaffold for PDE4, we analyzed the cocrystal structure of this scaffold candidate and PDE4D, and identified three potential sites of substitution based on the potential to make additional favorable chemical interactions in the available space at the active site. Consequently, a small set of compounds with substitutions at the 1-, 3- or 5-positions of the pyrazole were synthesized or purchased. We tested these compounds in the PDE assay, and then measured the IC_{50} values of these pyrazoles against a full panel of PDEs (**Supplementary Table 1**). The deletion of the ethyl group from the carboxylate substantially increased the IC_{50} from 60 μM to greater than 200 μM against PDE4B (**Table 1**). This is because the binding of the pyrazole carboxylate positions the ethyl group in the hydrophobic Q_1-pocket (**Fig. 1**), and deletion of the ethyl group eliminates the favorable hydrophobic interactions and therefore decreases the inhibitor's potency. Comparison of the methyl group
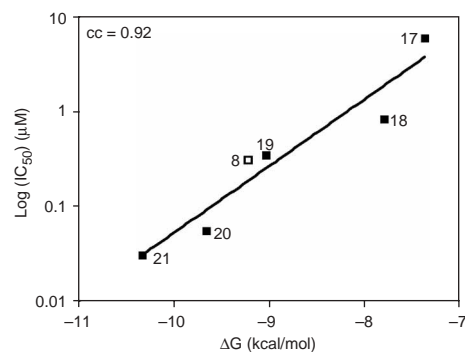

Figure 2 Correlation of calculated binding energy with experimental IC_{50} against PDE4B. The open square represents the derivative scaffold, 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 8), for which cocrystal structure was available at the time of the calculation. The complex structures for other compounds (represented by solid squares) were modeled using 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 8) as the template. The scaffold and designed compounds were subject to 500 ps molecular dynamics simulation, followed by MM/PBSA binding energy calculation. The IC_{50}s for the derivative compounds were obtained after they were synthesized.
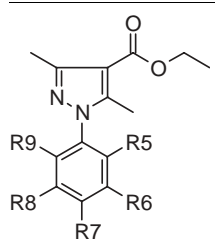
at the R2 and R3 positions with various substitutions of similar size, such as hydrogen, amine and trifluoromethyl, revealed that the methyl group at both of these positions is preferable (**Table 1**). Comparison of various aryl substitutions at R1, R2 and R3 positions indicated that the phenyl substitution at the R1 position is most potent. Therefore, we selected 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (PhPCEE, pyrazole no. 8, MW = 244) as a derivative scaffold for further optimization of potency.

We determined the cocrystal structure of PhPCEE in complex with PDE4D. This cocrystal structure revealed that the phenyl substitution at the 1-position of the pyrazole ring does not change the binding mode of the pyrazole in PDE4, thus validating the PCEE moiety as a scaffold for PDE4 (**Fig. 1b**). Furthermore, the 1-phenyl substitution gains several additional hydrophobic interactions with the conserved residues H160[4D], M273[4D] and L319[4D] in the active-site pocket. The phenyl ring binds at the deeper end of the pocket and points towards the di-metal ions. These interactions resulted in the substantially increased potency of this compound in PDE4B (IC_{50} = 270 nM) as compared with its parent compound (IC_{50} = 60 μM).

The structure of PCEE revealed that the ethyl group of the PCEE scaffold forms favorable hydrophobic interactions with a subpocket of the active site (**Fig. 1**). The deletion of the ethyl group from the carboxylate eliminated the activity, suggesting that the ethyl group is an integral part of the scaffold.

The third step: chemical optimization

The third step in our scaffold-based drug design approach is to optimize the validated scaffold into a potent and selective inhibitor. Derivative compounds based on the scaffold are designed using the cocrystal structure of the scaffold as a guide. Virtual libraries are computationally enumerated using available reagents according to the synthetic schema. The compounds' binding positions are predicted using scaffold-anchored docking and refined by molecular dynamics simulations, yielding a more detailed model of binding that includes an explicit contribution from solvent water. The compounds are then scored, and ranked according to their predicted ability to achieve potency by forming favorable interactions with

Table 2 SAR of substitutions on the phenyl group of the 1-phenyl carboxyl pyrazole derivative scaffold

R5	R6	R7	R8	R9	IC$_{50}$ (μM; PDE4B)	IC$_{50}$ (μM; PDE4D)
H-	CF$_3$-	H-	CF$_3$-	H-	160	97
NO$_2$-	H-	NO$_2$-	H-	H-	17	4.6
H-	H-	CF$_3$O-	H-	H-	32	21
H-	H-	(CH$_3$)$_3$C-	H-	H-	19	14
H-	H-	NH$_2$SO$_2$-	H-	H-	13	9.1
H-	H-	CH$_3$O-	H-	H-	5.9	2.0
H-	H-	NO$_2$-	H-	H-	0.84	1.0
H-	H-	NH$_2$-	H-	H-	0.35	0.16
Cl-	H-	H-	H-	H-	0.056	0.019
H-	NO$_2$-	H-	H-	H-	0.033	0.021

IC$_{50}$ values are in μM.

active site residues of the target protein. The availability of the scaffold structure greatly enhances the accuracy and predictability of structure-guided design, and the combination of computational prediction and chemistry insight sharply reduces the number of compounds that must be synthesized.

Because of the improved potency and the conserved binding mode of the PCEE scaffold and PhPCEE (**Fig. 1a,b**), we chose PhPCEE as the new starting point for lead optimization. We predicted that derivatives with small substitutions on the 1-phenyl ring were likely to retain the same mode of binding for the pyrazole ring and to be more potent. Drawing on the available reagent pool, we designed more than 100 compounds *in silico*. These compounds were docked into the PDE4B active site pocket with the scaffold part anchored at the observed binding position. The binding energy was also estimated using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method[22,23]. After eliminating a large number of substitutions that were predicted to cause undesirable interactions with residues in the active site, we synthesized ten compounds predicted to have increased binding affinity due to the formation of favorable interactions. The IC$_{50}$ values of these pyrazole derivatives (pyrazoles nos. 12–21) were measured (**Supplementary Table 1**). The calculated binding energy was similar to the experimentally determined IC$_{50}$ values, with an overall correlation of 0.92 between the predicted and measured values (**Fig. 2**). The high correlation between the predicted binding energy and the experimental IC$_{50}$ values suggests that these pyrazoles bind the active site of PDE4 in a similar manner, with the pyrazole ring adopting a conserved binding mode.

We determined the cocrystal structures of four of the ten synthesized molecules (**Fig. 1c–f**): two with increased potency (pyrazole nos. 20 and 21, **Table 2**), one with decreased potency (pyrazole no. 17) and one with similar potency (pyrazole no. 19) as compared with the parent compound PhPCEE. The pyrazole with a NO$_2$-substitution at the meta-position of the phenyl ring (pyrazole no. 21, MW = 289 Da) formed H-bonds with T345[4B], D392[4B] and the two water molecules coordinating the Zn^{2+}. The IC$_{50}$ of this compound decreased to 33 nM from 310 nM for its parent compound. Note that this compound
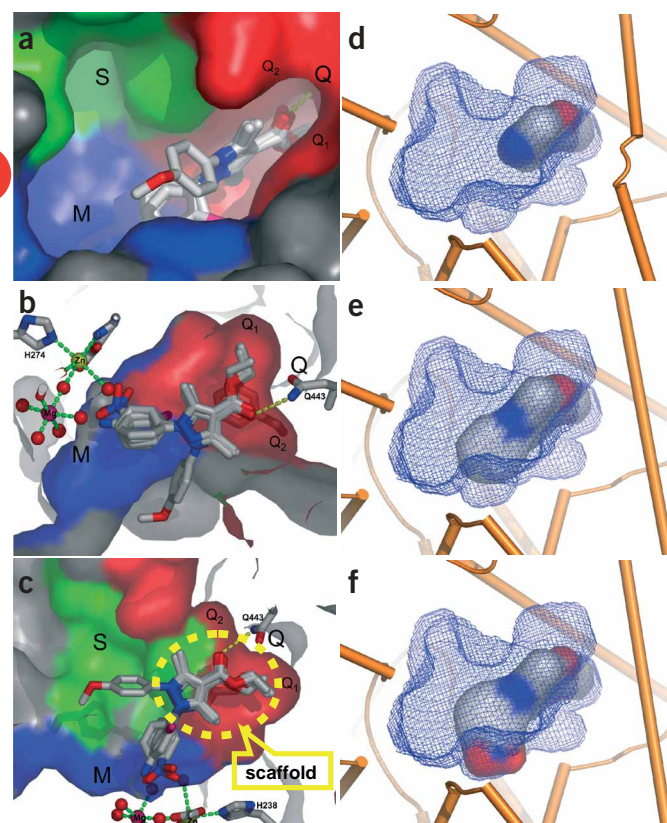

Figure 3 Pyrazole scaffold bound to PDE4B and PDE4D and the discovery of potent pyrazole inhibitors for PDE4 in three steps. Superposition of six different pyrazoles (nos. 2, 8, 17, 19, 20 and 21) in seven cocrystal structures with PDE4B and/or PDE4D revealed the consistent binding mode of the scaffold moiety (panels **a–c**). For clarity, only several side chains for one PDE4B cocrystal structure are shown. The three pockets in the active site are highlighted on the solvent accessible surface: the metal binding pocket (M) in blue, purine-selective glutamine and hydrophobic clamp pocket (Q) in red (which is further divided into Q$_1$, Q$_2$ sub-pockets) and solvent-filled side pocket (S) in green. The discovery of potent pyrazole inhibitors for PDE4 in three steps is illustrated in panels **d–f**. (**a**) A view looking down into the active site. The pyrazole carboxylate scaffold fits into the narrow passage formed by the hydrophobic clamp. (**b**) A view looking away from the S pocket. The pyrazole carboxylate scaffold forms an H-bond with the invariant Q443[4B]. (**c**) A view looking towards the S pocket. The ethoxy group occupies the Q$_1$-pocket. The scaffold that the six different pyrazoles share is marked by a dashed oval. (**d**) Scaffold discovery. The scaffold candidate, 3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 2), is a weak PDE4D inhibitor with IC$_{50}$ of 82 μM. (**e**) Scaffold validation. The derivative of the scaffold, 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 8) has significantly increased potency towards PDE4D with IC$_{50}$ of 0.27 μM. (**f**) Chemical optimization. The validated scaffold was optimized into a potent PDE4D inhibitor, 3,5-dimethyl-1-(3-nitro-phenyl)-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 21), with IC$_{50}$ of 0.021 μM. A 4,000-fold increase in potency was achieved in two rounds of chemical synthesis. Compounds are represented by solid surface colored by atomic types. The active site is represented by the blue mesh. The PDE4D is represented by cartoons where helices are shown as cylinders and loops are shown as tubes.

Protein	Compound #	Structure	D (Å)	C (%)	R_{sym}	M	R	R_{free}	RMSD Bond (Å)	RMSD Angle (°)	RMSD Torsion (°)	Rama-C (%)	Rama-A (%)	Rama-G (%)	Rama-D (%)
PDE4D	2		1.4	98	0.048	6.8	0.175	0.185	0.006	1.064	2.659	93.60	6.40	0.00	0.00
PDE4D	8		1.67	99.6	0.119	3.7	0.177	0.19	0.007	1.096	2.745	93.20	6.80	0.00	0.00
PDE4D	17		1.7	99.3	0.066	2.9	0.164	0.189	0.01	1.269	3.076	93.60	5.80	0.60	0.00
PDE4D	19		2.1	98.7	0.173	3	0.188	0.217	0.01	1.203	2.725	92.80	7.10	0.10	0.00
PDE4B	20		2.4	100	0.072	4.6	0.245	0.299	0.01	1.198	2.583	87.90	11.30	0.80	0.00
PDE4B	21		2.55	87.3	0.11	4.1	0.241	0.295	0.014	1.468	2.854	88.40	10.90	0.70	0.00
PDE4D	21		1.36	90.1	0.05	4	0.175	0.187	0.006	1.062	2.717	93.40	6.40	0.20	0.00

Figure 4 Data collection, processing, and refinement statistics for all the cocrystal structures of various pyrazoles in complex with PDE4B or PDE4D. The cocrystals of PDE4B with compounds grew in the $P2_12_12_1$ space group and have average cell dimensions of a = 89 Å, b = 94 Å and c = 107 Å. The crystals of PDE4D with compounds grew in the $P2_12_12_1$ space group and have average cell dimensions of a = 60 Å, b = 80 Å and c = 164 Å. Symbols used in the table are: D, diffraction resolution; C, completeness; M, multiplicity; R, crystallographic R-value; R_{free}, free R value (with 5% reflections in the cross validation data set); Rama-C, core region of Ramachandran plot; Rama-A, allowed region of Ramachandran plot; Rama-G, generously allowed region of Ramachandran plot; Rama-D, disallowed region of Ramachandran plot.

the di-metal ions to avoid the repulsive interactions (**Fig. 1f**). Accordingly, this compound has lost binding affinity by about tenfold compared with its parent compound, which has an IC_{50} of 5.9 µM (**Table 2**).

DISCUSSION

We have demonstrated the power of our scaffold-based drug discovery approach through the discovery and optimization of the pyrazoles. The binding mode of the pyrazole carboxylic ester scaffold does not change when substituents are added (**Fig. 3a–c**). This predictable binding mode, anchored by the pyrazole carboxylic ester scaffold, greatly increased the efficiency of chemical optimization. In total, a 4,000-fold increase in potency was achieved between the initial pyrazole carboxylic ester scaffold (IC_{50} = 82 µM in PDE4D) and the 2-chlorophenyl and 3-nitrophenyl pyrazoles (IC_{50}, 20 nM in PDE4D) in two rounds of chemical synthesis (**Fig. 3d–f**); only 21 compounds had to be derived. The total molecular weight added to attain this 4,000-fold increase in potency was only 121 Da. Because the most potent compounds are still relatively small, additional substitutions can be made to further exploit available areas of the active site to achieve higher potency and selectivity.

Our scaffold-based drug discovery approach bears some similarity to the fragment-based drug discovery approaches[2–6]. They all use biophysical techniques as the primary screen against a small library of low-molecular-weight compounds. The fragment-based methods use a small library of low-molecular-weight (<150 Da) chemical building blocks (about 1,000). These low-molecular-weight fragments generally bind target proteins very weakly (in the mM range) and cannot be detected by conventional high-throughput screening methods. Consequently, the fragment-based approaches use biophysical techniques, such as NMR[4] or X-ray crystallography[2,3,5,6], as primary screening methods. The low-molecular-weight fragments generate a small compound library, which is especially suited for screening by these biophysical techniques. These methods are very efficient[5], as a single preformed crystal can be soaked in a cocktail of 8–10—and sometimes even up to 100—compounds.

There are a number of unique features that differentiate our scaffold-based drug discovery approach from the fragment-based drug discovery approaches[2–6]. First, the scaffold library includes compounds that are significantly larger than basic building blocks, with an average molecular weight of about 250 Da. This increases the size of the scaffold library to about 20,000 compounds. However, the scaffold-like compounds contain functional groups that enable them to bind to the protein target at an affinity that is detected by high-throughput screening methods. Second, our scaffold-based drug discovery approach uses a low-affinity biochemical high-throughput screening at a high compound concentration of 200 µM for the initial selection of scaffold candidates. Third, the target protein as well as several other members in the protein family are screened,

binds to PDE4B and PDE4D with almost equal potency (**Table 2**) and in a very similar binding mode (**Fig. 1c**). Similarly, the Cl-substitution (pyrazole no. 20, MW = 278 Da) at the ortho-position makes several hydrophobic contacts with residues M347[4B], L393[4B] and F446[4B] (**Fig. 1d** and **Table 2**). These extra hydrophobic interactions increased the binding affinity about tenfold relative to its parent compound to an IC_{50} of 56 nM (**Table 2**). In contrast, a substitution at the para-position by a NH_2 group (pyrazole no. 19, MW = 259 Da) hardly improved the binding affinity (IC_{50} of 350 nM compared with an IC_{50} of 310 nM for its parent compound; **Table 2**). This para-substitution positioned the NH_2 group close to and pointing at the di-metal ions (**Fig. 1e**). This amine group formed three H-bonds with three water molecules, two of which are coordinated to the Mg^{2+}. However, this amine nitrogen is in close proximity to the carbon atom in M273[4D], which is unfavorable. The net result of these favorable and unfavorable interactions is a small decrease in binding affinity. The substitution of an OCH_3 group (pyrazole no. 17, MW = 274 Da) positioned this methoxy group too close to the di-metal ions and created some repulsion with the di-metal ions. This caused the entire methoxy phenyl pyrazole group to rotate 180° along the bond between the pyrazole carbon and the carboxylate carbon. Consequently, this methoxy phenyl group points away from

which substantially reduces the number of false positives. Only compounds that inhibit multiple members of the target family at low affinity will be selected for further screening by cocrystallography. Finally, the scaffold validation step ensures that the binding mode is tolerant to substitution, which leads to more predictable structure-activity relationships and improves the efficiency of chemical optimization.

We have developed a strategy that enables rapid and efficient scaffold-based design of inhibitors for a given target protein. We have applied our scaffold-based drug discovery platform to rapidly discover the pyrazole carboxylic ester scaffold and to optimize it into potent PDE4 inhibitors. The pyrazole carboxylic ester scaffold was discovered based on its two key interactions with PDE4 as observed in the cocrystal structure: the pyrazole ring is sandwiched by the hydrophobic clamp and the carboxyl oxygen forms a hydrogen bond with the purine-selective glutamine. This pyrazole carboxylic ester scaffold was validated through analysis of a small library of compounds with substituents on the pyrazole ring and the subsequent cocrystal structure of the phenyl-substituted compounds with PDE4. Another small library of compounds with substituents on the phenyl ring of the derivative scaffold resulted in the discovery of inhibitors with low nanomolar potency. The robustness and efficiency of the scaffold-based drug discovery method should make it widely applicable to expedite the lead discovery effort for many other targets for which known small molecule modulators are limited.

METHODS

The scaffold-library design. The scaffold-library was prepared by analyzing the available compounds from 17 different chemical vendors for compounds in the molecular weight range of 120–350 Da. In the first step, we chose 1,994,133 compounds for analysis based on their molecular weight and discarded those compounds with reactive groups; this gave us 275,555 compounds in the desired molecular weight range. These 275,555 compounds were broken down into potential scaffold components by dividing each compound into smaller substructures, through a process of fragmenting at rotatable bonds within each molecule, resulting in a total of 1,277,373 individual substructures. These scaffold components were then grouped by chemical similarity, thus building a representation of all of the substructures within the starting 275,555 compounds. Further selection and processing to remove any compounds within a high similarity threshold (Tanimoto index > 0.85) of one another resulted in the final composition of the desired chemical library, which consisted of 20,360 compounds. These 20,360 compounds cover roughly 80% of the clustered scaffold component space, and compose the core 20 K scaffold library subsequently used for screening.

Cloning, protein expression and purification. The cDNA encoding the catalytic domains of human PDE2A (gene number: NM_002599; coding region: H574-E941), PDE3B (NM_000922; Q650-A1084), PDE5A (NM_001083; S531-N875), PDE7B (NM_018945; Q91-P450), PDE8A (NM_02605; M1-E829), PDE9A (NM_002606; S226-A593), PDE10A (NM_006661; M432-D779) and PDE11A (NM_016953; D633-N988) were cloned into pET15S vectors (Novagen) in which a His-tag is appended to the coding sequence. The cDNA encoding the full-length human PDE8A were cloned into pFastBac vector with an N-terminal His-tag. The cloning of catalytic domains of human PDE1B, PDE4B and PDE4D as well as the expression and purification of all the PDE catalytic domains were based on the protocols reported previously[24].

Assay of phosphodiesterase activity. Measurement of phosphodiesterase activity takes advantage of the selective binding of 5'-AMP or 5'-GMP (and not cAMP or cGMP) to yttrium silicate beads with embedded scintillant. Briefly, 0.1–1 nM PDE is incubated with 50 nM ^{3}H-cyclic nucleotide (Amersham, 5–60 Ci/mmol) in 50 mM Tris, pH 7.5, 8.3 mM $MgCl_2$, 1.7 mM EGTA, 0.01% BSA at 30 °C for 30 min in 384-well assay plates. The assay is terminated by adding one-third volume of 5 mg/ml yttrium silicate beads in 18 mM zinc acetate/zinc sulfate solution (3:1). A minimum of 30 min after mixing and centrifuging the reaction, hydrolysis is quantified by reading the plates in a scintillation counter (Trilux, Wallac).

Synthesis of pyrazoles. The pyrazoles were either purchased from commercial sources or synthesized by reacting the corresponding 1,3-dicarbonyl compounds or 2-acetyl-3-oxo-butyric acid ethyl-ester with hydrazines or phenyl hydrazines[25]. As an example, 3,5-dimethyl-1-(3-nitro phenyl)-1H-pyrazole-4-carboxylic acid ethyl ester (pyrazole no. 21) was synthesized by refluxing ethyl diacetylacetate with 3-nitro-phenylhydrazine in ethanol and acetic acid mixture in the presence of molecular sieves for 3 h, evaporating the solvent and purifying by column chromatography (silicagel, 1:9 ethylacetate:hexane as eluent). The compound was obtained as a yellow solid with an isolated yield of 94%.

Library enumeration and MM/PBSA-based scoring. Library enumerations were done with LEGION from the Sybyl program suite (Tripos Associates) with inputs selected based on the synthetic schema. The 3D structures of the virtual compounds were generated using the program CONCORD (Tripos Associates). The structures were then docked into the active site using a proprietary fitting/minimization algorithm relying on the crystal structure of the scaffold as the anchor.

The docked structures were used as the starting points for 500 ps molecular dynamics simulations using AMBER7[23] and PARM94 force field[26]. The atomic partial charges of compounds were generated semi-empirically using MOPAC[27]. The complex binding free energy (ΔG) was estimated using the MM/PBSA methodology[22] implemented in-house. Further details of the methodology will be discussed in a separate publication. Because of the difficulty in calculating the absolute entropic contribution to the binding affinity for such large molecular system, the MM/PBSA energies were normalized by a constant that allows the log(IC_{50})-ΔG regression line to pass through the (1 μM, −8.2 kcal/mol) point.

Protein crystallization, data collection and structure refinement. All proteins were crystallized using the sitting drop method. Crystals of PDE4B with various compounds were grown at 4 °C using the sitting drop method, by mixing equal volumes of the protein at 10 mg/ml with 1.6–2.1 M ammonium sulfate, buffered by 0.1 M 3-(cyclohexylamino)propane sulfonic acid (CAPS) to a pH range of 10.0–10.5, in the presence of 0.2 M lithium sulfate and 1 mM compound. Crystals of PDE4D with various compounds were grown at 15 °C using the sitting drop method, by mixing equal volumes of the protein at 30 mg/ml with a well buffer of 19–24% PEG 3350; 28–34% ethylene glycol; 10% isopropanol, buffered by 0.1 M Bis-Tris propane to a pH range of 6.0–8.5 and adding the compound to 1 mM concentration. Ranges of pH and precipitants and additives were used, to help increase the chances of cocrystallization, as the initial 316 compounds were from different chemical classes, and likely to have different effects on the crystallization mixture, e.g., pH, protein solubility. Overall, about 85% of the compounds have been successfully cocrystallized with their protein target, PDE4B or PDE4D.

X-ray diffraction data were collected at BL8.3.1 of the Advanced Light Source (Lawrence Berkeley National Laboratory, Berkeley, California) or at the Stanford Synchrotron Radiation Laboratory (Stanford, California). The data were processed by MOSFLM[28] and SCALA[29] driven by the ELVES[30] automation scripts. The cocrystal structures of PDE4B and PDE4D were solved by molecular replacement using EPMR with the PDE4B apo structure[31], the PDE4D apo structure[32]. The overall hit rate of each cocrystal data set yielding traceable electron density that corresponds to the compound is about 50% for PDE4B and about 30% for PDE4D. The cocrystal structures were refined by CNX[33] and REFMAC[34] with intermediate stages of manually rebuilding in O[35]. The relevant data collection and refinement statistics are given in **Figure 4**.

Coordinates. The coordinates and structure factors for the cocrystal structures of PDE4B and PDE4D with pyrazoles have been deposited in the Protein Data Bank with the accession numbers 1Y2B, 1Y2C, 1Y2D, 1Y2E, 1Y2H, 1Y2J and 1Y2K.

Note: Supplementary information is available on the Nature Biotechnology website.

COMPETING INTERESTS STATEMENT
The authors declare competing financial interests (see the *Nature Biotechnology* website for details).

1. Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
2. Erlanson, D.A., McDowell, R.S. & O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **47**, 3463–3482 (2004).
3. Rees, D.C., Congreve, M., Murray, C.W. & Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discov.* **3**, 660–672 (2004).
4. Shuker, S.B., Hajduk, P.J., Meadows, R.P. & Fesik, S.W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
5. Nienaber, V.L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **18**, 1105–1108 (2000).
6. Erlanson, D.A., Wells, J.A. & Braisted, A.C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
7. Houslay, M.D. Adaptation in cyclic AMP signalling processes: a central role for cyclic AMP phosphodiesterases. *Semin. Cell Dev. Biol.* **9**, 161–167 (1998).
8. Beavo, J.A. Cyclic nucleotide phosphodiesterases: functional implications of multiple isoforms. *Physiol. Rev.* **75**, 725–748 (1995).
9. Conti, M. & Jin, S.L. The molecular biology of cyclic nucleotide phosphodiesterases. *Prog. Nucleic Acid Res. Mol. Biol.* **63**, 1–38 (1999).
10. Francis, S.H., Turko, I.V. & Corbin, J.D. Cyclic nucleotide phosphodiesterases: relating structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* **65**, 1–52 (2001).
11. Corbin, J.D. & Francis, S.H. Pharmacology of phosphodiesterase-5 inhibitors. *Int. J. Clin. Pract.* **56**, 453–459 (2002).
12. Rotella, D.P. Phosphodiesterase 5 inhibitors: current status and potential applications. *Nat. Rev. Drug Discov.* **1**, 674–682 (2002).
13. Souness, J.E., Aldous, D. & Sargent, C. Immunosuppressive and anti-inflammatory effects of cyclic AMP phosphodiesterase (PDE) type 4 inhibitors. *Immunopharmacology* **47**, 127–162 (2000).
14. Mehats, C., Andersen, C.B., Filopanti, M., Jin, S.L. & Conti, M. Cyclic nucleotide phosphodiesterases and their role in endocrine cell signaling. *Trends Endocrinol. Metab.* **13**, 29–35 (2002).
15. Beavo, J.A. & Brunton, L.L. Cyclic nucleotide research–still expanding after half a century. *Nat. Rev. Mol. Cell Biol.* **3**, 710–718 (2002).
16. Conti, M. Phosphodiesterases and cyclic nucleotide signaling in endocrine cells. *Mol. Endocrinol.* **14**, 1317–1327 (2000).
17. Conti, M. *et al.* Cyclic AMP-specific PDE4 phosphodiesterases as critical components of cyclic AMP signaling. *J. Biol. Chem.* **278**, 5493–5496 (2003).
18. Houslay, M.D. & Adams, D.R. PDE4 cAMP phosphodiesterases: modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization. *Biochem. J.* **370**, 1–18 (2003).
19. Jin, S.L. & Conti, M. Induction of the cyclic nucleotide phosphodiesterase PDE4B is essential for LPS-activated TNF-alpha responses. *Proc. Natl. Acad. Sci. USA* **99**, 7628–7633 (2002).
20. Bardelle, C. *et al.* Phosphodiesterase 4 conformers: preparation of recombinant enzymes and assay for inhibitors. *Anal. Biochem.* **275**, 148–155 (1999).
21. Card, G.L. *et al.* Structural basis for the activity of drugs that inhibit phosphodiesterases. *Structure* **12**, 2233–2247 (2004).
22. Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A. & Case, D.A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **120**, 9401–9409 (1998).
23. Case, D.A. *et al. AMBER 7* (University of California, San Francisco, 2002).
24. Zhang, K.Y.J. *et al.* A glutamine switch mechanism for nucleotide selectivity by phosphodiesterases. *Mol. Cell* **15**, 279–286 (2004).
25. Richter, W. 1,3-Cyclopentanedione and isomeric enol-lactones. IV. Some reactions of diacetic acid ethyl ester. *Helv. Chim. Acta* **35**, 478–485 (1952).
26. Cornell, W.D. *et al.* A second generation force field for the simulation of proteins, nucleic acid, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
27. Stewart, J.J.P. *MOPAC Manual* (United States Air Force Academy, Colorado Springs, CO, 1985).
28. Leslie, A.G. Integration of macromolecular diffraction data. *Acta Crystallogr.* **D55**, 1696–1702 (1999).
29. Evans, P.R. Data Reduction. *Proceedings of CCP4 Study Weekend, On Data Collection and Processing, Daresbury, Warrington, UK*, 114–122 (1993).
30. Holton, J. & Alber, T. From the cover: automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. USA* **101**, 1537–1542 (2004).
31. Xu, R.X. *et al.* Atomic structure of PDE4: insights into phosphodiesterase mechanism and specificity. *Science* **288**, 1822–1825 (2000).
32. Huai, Q. *et al.* Three-dimensional structures of PDE4D in complex with roliprams and implication on inhibitor selectivity. *Structure* **11**, 865–873 (2003).
33. Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905–921 (1998).
34. Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. & Dodson, E.J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr.* **D55**, 247–255 (1999).
35. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110–119 (1991).